

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2003, 13:02:56 ; Search time 29 seconds
(without alignments)
1191.372 Million cell updates/sec

Title: US-10-027-000-2

Sequence: 1 MADIDVEALIKRLTAEKVD.....DSVALRGKFTVGETYMGSGV 833

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Swissprot_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2052	46.7	845	1	BGLS_KLUMA
2	1445.5	32.9	818	1	BGLS_AGRU
3	1145	26.1	754	1	BGLB_CLOTH
4	722	16.4	880	1	P22507 saccharomyc
5	706	16.1	876	1	BGLI_SACFI
6	646.5	14.7	765	1	BGLX_SALTY
7	635.5	14.5	765	1	BGLX_ECOLI
8	593	13.5	825	1	BGLS_HANAN
9	581	13.2	860	1	BGLI_ASPAC
10	451.5	10.3	830	1	BGLS_BUPFI
11	425	9.7	947	1	BGLS_RUMAL
12	219.5	5.0	642	1	YRBD_BACSU
13	164.5	3.7	192	1	BGLS_SCHCO
14	151.5	3.5	330	1	NAGZ_VIBCH
15	147.5	3.4	329	1	NAGZ_VIBBU
16	147	3.3	343	1	NAGZ_YERPE
17	144	3.3	341	1	NAGZ_ECOLI
18	140	3.2	341	1	NAGZ_PASU
19	139	3.2	351	1	NAGZ_PASU
20	136	3.1	341	1	NAGZ_SALTY
21	134	3.1	341	1	NAGZ_SALTY
22	133.5	3.0	332	1	NAGZ_PSEAE
23	130	3.0	568	1	DCPY_ZYMO
24	129	2.9	351	1	NAGZ_HAEIN
25	128.5	2.9	598	1	HEXA_ALTSO
26	125	2.8	1058	1	CARB_FUSUN
27	123.5	2.8	380	1	PEL_PSEMA
28	121.5	2.8	634	1	DHSA_ARATH
29	121.5	2.8	2647	1	FLNA_HUMAN
30	121	2.8	661	1	DHSA_DROME
31	120.5	2.7	664	1	DHSA_HUMAN
32	119.5	2.7	1070	1	AGU_CANTS
33	119	2.7	1953	1	BIGA_SALTY

34	119	2.7	2774	1	MAPA_RAT
35	118.5	2.7	599	1	IAC2_PNCU
36	118.5	2.7	1064	1	CARB_LACIA
37	117.5	2.7	1178	1	RPOC_CLOPE
38	116.5	2.7	1064	1	CARB_LACIC
39	116	2.6	914	1	BPBA_BACSU
40	116	2.6	1057	1	CARB_STAM
41	115.5	2.6	1058	1	CARB_LACPL
42	115	2.6	826	1	VILI_CHICK
43	115	2.6	1057	1	CARB_STAM
44	113.5	2.6	646	1	DHSA_CAEEL
45	113.5	2.6	1171	1	TRIZ_STRCO

ALIGNMENTS

RESULT 1
BGLS_KLUMA STANDARD: PRT; 845 AA.

AC BGLS_KLUMA
AC P07337;
AC 01-APR-1988 (rel. 07, Created)
AC 01-APR-1988 (rel. 07, Last sequence update)
AC 15-JUN-2002 (rel. 41, Last annotation update)
AC Beta-glucosidase precursor (EC 3.2.1.21) (Gentloblase) (Cellulobase)
AC (Beta-D-glucoside glucohydrolase).
AC Kluyveromyces marxianus (Yeast) (Kluyveromyces fragilis)
AC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
AC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
AC NCBI_Taxid=4911;
AC (1)
AC SEQUENCE FROM N.A.
AC STRAIN=ATCC 12424;
AC MEDLINE=88210533; PubMed=2835179;
AC Raynal A., Gerbaud C., Francinques M.C., Guerin M.,
AC Kluyveromyces fragilis cloned in Saccharomyces cerevisiae.";
AC Curr. Genet. 12:175-184(1987).
AC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
AC glucose residues with release of beta-D-glucose.
AC -1- PATHWAY: Cellulose degradation.
AC -1- SUBUNIT: HOMOTETRAMER.
AC -1- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES.
AC
AC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
AC between the Swiss Institute of Bioinformatics and the EMBL outstation -
AC the European Bioinformatics Institute. There are no restrictions on its
AC use by non-profit institutions as long as its content is in no way
AC modified and this statement is not removed. Usage by and for commercial
AC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
AC or send an email to license@isb-sib.ch).
AC EMBL: X05918; CAA29353.1; .
AC PIR: A29148; GLVK.
AC InterPro: IPR002772; GH_3C.
AC InterPro: IPR001764; GH_3N.
AC Pfam: PF00933; Glyco_hydro_3_1.
AC Pfam: PF01915; Glyco_hydro_3_C_1.
AC PRINTS: PR00133; GLYDRLASE3.
AC PROSITE: PS00775; GLYCOSYL_HYDROL_F3; 1.
AC Hydroxylase; Glycosidase; Cellulose degradation; glycoprotein; Signal.
AC SIGNAL
AC CHAIN
AC ACT_SITE 225 845
AC CARBOHYD 66 66
AC CARBOHYD 304 304
AC CARBOHYD 438 438
AC CARBOHYD 621 621
AC SEQUENCE 845 AA; 93916 MM; 78DAC1550D592B5 CRC64;
AC
AC BEST LOCAL SIMILARITY 46.7%; Score 2052; DB 1; Length 845;
AC Query Match 47.2%; Pred. No. 5.9e-125;
AC Matches 403; Conservative 152; Mismatches 269; Indels 30; Gaps 11;

Qy	1	MADIDEALIKKTLIAEKVDLAGIDFHWTKLPRHCVPSLFITDGPNGVRCYKFNCP	60
Db	1	MSFEDVQULSELSEINODEKSTLISAVDFMHTKIERIDJIPAVNSDGPNGINGITFFDVP	60
Qy	61	ACFPQSTISLGSFNFOTLTLEAGCKMMGCEAIKSAHVILIGPTINNORSPLGGRSEIGE	120
Db	61	SCGFPGTGLASFDDRLDLETAGKILMAKESIAKNAAVILGPTTNQRPLOGRPFESSE	120
Qy	121	DPLIABLGAALIRIGIQSTGVQATIKHFLCNDQEDRRMWSIYTERALREIYALPOIA	180
Db	121	DPLIAMAAVSVVKGQOGGIAATYKHFVCHDLEDFORSNSIYSERLIREIYEPRLA	180
Qy	181	VRDSQGAFTMYNGANGCSEENKYLDGMLRKQMGDGLIMSDWGTQSTETAVYAGL	240
Db	181	VKIANVVCIMTYANVKNYNGHCSQSKLLIDILIRDKMKDGLIMSDWGTQSTETAAIINGL	240
Qy	241	DLEMGPFRFKEETLKFNYSNGKPIHV-IDORAREVLOFVK---KCAASVTENGPE	294
Db	241	DIEFPGRFMRTRALVSHLSINREQITTEEDVDHRYQLKMKIYVONLEKGTIVENGPE	300
Qy	295	TYVNNPPEALALRRVNGDVIYLLKNENNVPLSKKKKTLIVGNAAQATYHGGCSAALR	354
Db	301	STSNNTKESYDLRIEADSTIYLLKNKNNYLTSEKEROYHVIIGNAKAKISSGGSSASN	360
Qy	355	AYYAVTPEDGSLKOLETPSYTCAGATYVP-PIEGEOLFPDPAZ---GMMWRJENE	407
Db	361	SYVYVSPREGIYINKLGKEVDYTVGAYSHKISIGGLAESLIDAKPRADAENAGLIAKYSN	420
Qy	408	PQGTNRQHOHIDELF---FYTDMHLVDYVYHPRKADT---WYADMEGTYTADEDCYIELG	460
Db	421	P---VEERSEDEPFFHVTKVNSNHLPEFKHEKVDPKNPFEFVTLTGQYVPOEDGDIYFS	478
Qy	461	LYVCGTAKAYVDDOLVYVNAKROYPGDAFPCSATREEGRNLYKGMTYAFKLEFGSAPR	520
Db	479	LQVYSSGLIEYINDELIDQKRNQERGSCTFAGIKERTKULILKKGQVYVWRYEGSGPT	538
Qy	521	YTLKADTVPGHGLIRNGGCKVIDDQAEIERKSVALAKEHDOVILICAGIINDMETEGADRA	580
Db	539	SOIWEF---GAGGFOAQVIAIADDEDEIRAAELAKHDKAVILIGLNGEMETEGDRE	595
Qy	581	SMKLEPQYDLOLIADYAAANPMTVVMVQGTPEEMPMIDATPAYOMYGGNETGNSIADY	640
Db	596	NMDLPKRNELVRAVLIKANPITVYVNOSSGPVPEPWLEEBNALVQMYGGNETGNAIADY	655
Qy	641	VFGDYNPBGKUSLSEPKRLQDNPAFLNFRTEAGRTIYGEDVYVGYRYFERADKDVPFPG	700
Db	656	LYGDVVPNGKLSLWMPKRLQDNPAFLNFKTEFGVYVGEDIFGVGYRYEKLQKKAFFPG	715
Qy	701	HGLSTYFAFNLVSXHKDGLKLSYSVKNTG--SVPAQVQAOIYVPRDIAAKINRYAKEL	759
Db	716	YGLSTTFIELDISPFKATYDDKIDISVDYKNTGOKFAGSEVVOYVYFSALN-SKVSRYAKEL	774
Qy	760	KGFAKVELQPGETRAVYITEDQEKVYAVYFDEERDQMCVEKGDEYIVYSASAQKDVALAR	819
Db	775	KGFEKVHLEPGEKTYVNIETELKDAISYFNEELGKMHVVEAGELVSVGSS---DDILSY	831
Qy	820	GKFTYGEIYMMWSGV	833
Db	832	KEFVVKDLYWKGL	845

RESULT 2			
BGLS_AGRТУ			
ID	BGLS_AGRТУ	STANDARD;	PRT; 818 AA.
AC	P27034;		
DT	01-AUG-1992 (Rel. 23, Created)		
DT	01-AUG-1992 (Rel. 23, Last sequence update)		
DT	01-DEC-1992 (Rel. 24, Last annotation update)		
DE	Beta-glucosidase (EC 3.2.1.21) (Gentibiase) (cellulobase) (Beta-D-glucoside gluconhydrolase).		
GN	CBG-1.		
OS	Agrobacterium tumefaciens.		

OC	Bacteria	Proteobacteria:	alpha subdivision:	Rhizobiaceae group;
OC	Rhizobiaceae:	Rhizobium.		
OX	NCBI_TaxID=358;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=B3/73;			
RX	MEDLINE=92165721; PubMed=1537792;			
RA	Castle L.A., Smith K.D., Morris R.O.;			
RT	"Cloning and sequencing of an Agrobacterium tumefaciens			
RT	bet-glucosidase gene involved in modifying a vir-inducing plant			
RT	signal molecule."			
RL	J. Bacteriol. 174:1478-1486(1992).			
CC	-1- FUNCTION: INVOLVED IN MODIFYING A VIR-INDUCING PLANT SIGNAL			
CC	MOLECULE. HYDROLYZES CONIFERIN BUT NOT CELLOBIOSE.			
CC	-1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-			
CC	-1- SUGAR RESIDUES WITH RELEASE OF beta-D-glucose.			
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).			
CC	-1- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES.			
CC	-----			
CC	THE SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	at the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement. See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch .			
CC	-----			
CC	EMBL: M59852; AAA22082.1; .			
DR	PIR: A42292; A42292.			
DR	InterPro: IPR002772; GH_3C.			
DR	InterPro: IPR001764; GH_3N.			
DR	Pfam: PF00933; Glyco_hydro_3; 1.			
DR	Pfam: PF01915; Glyco_hydro_3_C; 1.			
DR	PRINTS: PR00133; GLHYDRLASE3.			
DR	PROSITE: PS00775; GLYCOSYL_HNDROL_F3; 1.			
FW	Hydrolase; Glycosidase			
FT	ACT_SITE 222			
FT	BY SIMILARITY			
SO	SEQUENCE 818 AA; 88290 MW; DOI6C8F02FEFF830 CRC64;			

Query Match:	32.9%	Score 1445.5	DB 1	Length 818
Best Local Similarity	40.5%	Pred. No. 8.8e-86		
Matches 334	Conservative 134	Mismatches 310	Indels 47	Gaps
QY	6	VEALLKLTLLAEKVDLLAGIDFMWTKALPKHGVPSLFTDGPNGVR-GTFKFNCPVAPACF	64	
Db	2	IIDIIIDKMTLEEQVSYLISGADPFWTTVAIERGVDPKIKVTIDGPRGAGAGGSLGCKXSACF	61	
QY	65	PGCSTISGTFNQTLLLEAGKMKMGKEALAKSAHHYILGPTINMORSPLGGRGFSICGEDPFL	122	
Db	62	PAVALIAGAWDELLERGVAGLGOAKSKASVYLAPTVAINHSGINGNFECYSEDPAL	121	
QY	125	AGLGAALLRGIOSTGYOATIKHFELCNDODRRMMYOSITERALREIYALPFOYAVDS	184	
Db	122	TAAACAAVYINGVQSGVAATIKHFVAAHSELEIQTMSSDYDERTILREIYLPFEEDAKKA	181	
QY	185	QPGAFMTAYNGINGVSCSENNKYLDGMLRKREWCMDGLINSDWYGTYSTTEAVVAGLDLDM	244	
Db	182	GVAAYVSSYNNKLINGTYSSENFMLLTJKVLREEMGFQDGVWDMFQSGHSTAETIINGDLEM	241	
QY	245	PGPPRRRGETLKFVNSNGKPRPIHHYIDRAEVLQFPAKKAASGVTEENGPETIVN--NTPE	302	
Db	242	PGPMWDRGKILVAAARECKYAAETVRAASARILLLEERY--GAEEKAPDALEAHALDPE	298	
QY	303	TALLRKVNGEGVILIKENNVLPLSKRK--KTIIYGPAAKQATYNGGSAALRAYAVT	360	
Db	299	DRAILLROIQARGAVULKND-GYLPPLAKSSPDQIAYGPRPAASARMGGSARRIAAHYTVS	357	
QY	361	PRDGSKOLETPS--YTVGAYTT--VPRILEOCITPRGAPGMKRVFNRPCTPRROH	416	
Db	358	PLEGIRALISNANSILRHAVCGCNNRLLIDVFSG--MTVEYFG--RGFESRPVHVEVE	412	
QY	417	IDELFF--TETDHHLVDDYHNPKAADLTWADMEGTYTADDEOSTYELGLVSGTAKAYVD	473	

Db 413 KGEFFWEDLPESGDLADLAF-----SARMTATFVPOETGEHIFGMTNAGIARLFLYVDG 463

QY 474 QLVVNDATKOVPPDAPFGSRTRETRGRINLVKNTYKKEIEGSAFTYLLKDDTYPGHG 533

Db 464 ELVVDYDGMWTKENFEFGANSORAVTLGAARRRVAYEX-EAPKASLDGINIC----- 518

QY 534 SLRVGGCKVDDOAEIEKSVALKAKEDQVYICGLNADWETEGADARSKKLPGLVDOLIA 593

Db 519 ALRFGVEKPLGD-AGTAENAVETARKSDIYLLVLRGSEMDTEGLDIPMRILPGROBELIE 577

QY 594 DVAANNTVYVWGTGPEEMPLDAPVAYIQAWYGNETGNSIADVFGDYNPSGKLSL 653

Db 578 AVAETNPVVVVLGTGPIEMWLGKRVAVLQWMPGDELGNALADVLFGDVEPAGRLPQ 637

QY 654 SFPRKLDQDNAPFLN----FTEGRLYGEDYVGYRYEAFDKNVPFGHLSYTFFA 709

Db 638 TFRKALDNLSDITDDPSIYFGDGHVRYAGIVGGRHHTREIELEFPFGELGYTRFT 697

QY 710 FSNLSYS-----HKDKLSYSLSVKNTGSVGAQVADLYVPLQAKINPVEKELCFANY 765

Db 698 WGAPOLSGTEMGADG-LTVYVDVTNIGDRAGSDVYQLYHS-PNAVREPRFKELRAFAKL 755

QY 766 ELQGETKAVTIEOEKYVAAYDEERDQWCVKGEYVSDS 810

Db 756 KLAGATGATVAKIAPRDL-AYFDVEAGRERADAGKVELIYAASA 799

RESULT 3

BGLB_CLOTM STANDARD; PRT; 754 AA.

AC BGLB_CLOTM

DT 01-JAN-1990 (Rel. 13, Created)

DT 01-JAN-1990 (Rel. 13, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Thermoslabile beta-glucosidase B (EC 3.2.1.21) (gentiobiase)

DE (Cellulobase) (Beta-D-glucoside glucohydrolase).

GN BGLB.

OS Clostridium thermocellum.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium

OX NCBI_TaxID=1515;

RN [1]

RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.

RC STRAIN=ATCC 27405 / DSM 1237;

RX MEDLINE=89364694; PubMed=2505054;

RA Graebnitz F., Ruecknagel K.P., Seiss M., Staudenbauer W.L.;

RT "Nucleotide sequence of the Clostridium thermocellum bgib gene

encoding thermostable beta-glucosidase B: homology to fungal

beta-glucosidases";

RT Mol. Gen. Genet. 217:70-76(1989).

-1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-

glucose residues with release of beta-D-glucose.

-1- PATHWAY: Cellulose degradation

-1- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation-

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC EMBL: X15644; CAA33665.1; -

DR PIR: S04381; S04381.

DR InterPro: IPR002772; GH_3C.

DR InterPro: IPR001764; GH_3N.

DR Pfam: PF00933; Glyco_hydro_3; 1.

DR Pfam: PF01915; Glyco_hydro_3_C; 1.

DR PRINTS: PR00133; GLHYDLASE3.

DR PROSITE: PS00775; GLYCOSYL-HYDROL_F3; 1.

KM Hydrolase; Glycosidase; Cellulose degradation.

FT ACT_SITE 231 231 BY SIMILARITY.

SO SEQUENCE 754 AA; 84100 MW; ABCEB38BFC9PD3A8 CRC64;

Query Match 26.1%; Score 1145; DB 1; Length 754;

Best Local Similarity 31.9%; Pred. No. 2,1e-66;

Matches 270; Conservative 132; Mismatches 249; Indels 196; Gaps 15;

QY 4 IDVEALIKRLTAEKVYDLAIDEMHTKALPKHGVPSLRFTDGPNGVGRGTA-----FF 56

Db 3 VDIKTIKQMTLEEKAGLCSGLDFWHTKRPVERLIGIPSIIMTDGPHGLKREDEAEIADIN 62

QY 57 NCVPACFPCCGSLSTGNTNLTLEAGKMKKEIAKSAHYILPTIMOSPSPGSGRFE 116

Db 63 NSVPATCEPSSAAGLACSDREIVERVYALGEBQCAENVSILLEGANIKSPLOGRFE 122

QY 117 SGEPPFLAGLGAALLINGIOSYGVATIKHPLCNDQEDRRMYOSIVTERALREIYALP 176

Db 123 YPEDEPILSSELAASHIKGVOSOGAGCLKHPAANNGHRRMTVOTIYDETKREIYFAS 182

QY 177 FOIAVRDSOPGAFMTAVNGINGVSCSENPKYLDGMLKREKMGDGLMSDWTGITSYTA 236

Db 183 FENAVKAKAPVVMCAVYNKNGEYSENRYLTLEVLKEMHMDGFVSDMGAVNDRVSGL 242

QY 237 VAGLDLEMPGPRFGEETLKEFVNSGKPIIHYIDQAREVLOFVKKCAASGVTEGPRETT 296

Db 243 DAGLDLEMTSGITDKKIVEAVKSGKLSENILNAVERILKIYIM-----ALEKKENA 297

QY 297 VANTPETAALLKRVNGEIVLLKNENNVLPISKKRKLTVGNPAQATYHGGSAALRAY 356

Db 298 QYEQDAHHRHLAQAAESVNLKNEEDVLPFKKSGTILAFGAKKPRGSGSSHI--- 354

QY 357 VAVTPEDGLSKOLETPPSYTVGATYTPPILEGOCITPDGAGMWRPNEPRGPNQH 416

Db 355 -----TPTR- 358

QY 417 IDELFETKDMHLVDYHHPKADTWYADMEGTFTADECTYELGLVCGTAKAYDDOLY 476

Db 359 LDDI-----YEELK-----KAGAD----- 372

QY 477 VDNATKQVGDGAFFGSATRETERGRINLVKNTYKKEIEGSAFTYLLKDDTYPGHSLR 536

Db 373 -----KVLVYSEGR-----LENDGI----- 389

QY 537 VGGCKVIDQAEIEKSVALKAKEDQVYICGLNADWETEGADARSKKLPGLVDOLIA 596

Db 390 -----DEELINEAKKKAASSDVAVPAAGLPDEESGCFRTHSTIPENORLIEAVY 441

QY 597 AANPNTVYVWGTGPEEMPLDAPVAYIQAWYGNETGN-----SIADVFGDYNPSGK 650

Db 442 EVGNSILVYVLLNSPYEMWIDKVKVLEAYLGQALGRWMCYSKSVI-----GK 494

QY 651 LSLSPKRLQDNAPFLNFTREAGRTLYGEDYVGYRYEAFDKNVPFGHLSYTFFA 710

Db 495 LAETFPVKLSHNSPYLNFGEEDRVEYKEGLFVGARYYDTKIEPLFPFGHLSYTFFA 554

QY 711 SNLSYSHKD-----GKLSYSLSVKNTGSVGAQVADLYVPLQAKINPVEKELCFANY 766

Db 555 SDISVKKKQVSDNSIINVSAYKNAVMKAGKELVQLYADVSS-VRRPEKELGPEKVF 613

QY 767 LQGETKAVTIEOEKYVAAYDEERDQWCVKGEYVSDS-----EVIYSDSAAKGYA 817

Db 614 LNGEKETVTF-TLDRARAFYVYTKDMHVEGSEGLILIGRSSRDIVLKESPVNSYVK 672

QY 818 LRGEFTV 824

Db 673 IRRKFTV 679

RESULT 4

BGL2_SACFI STANDARD; PRT; 880 AA.

ID BGL2_SACFI

DT 01-AUG-1991 (Rel. 19, Created)

DT 01-AUG-1991 (Rel. 19, Last sequence update)

	Query Match	16.4%	Score 722	DB 1:	length 880;
	Best Local Similarity	24.5%	Pred. No. 5,6e-39)		
	Matches 243;	Conservative 118;	Mismatches 267;	Indels 362;	Gaps 33;
QY	7	EALIKLITLAEKVYDLAIDF-----HTKALPKHGVDLSRLFTDGNVGRTGTFNGCV	59		
		: :			
Db	64	KALVSOMTIVKEKNLTGTG-WQLPCGVGGNSVPREGIIPNLCLIODPLGLRVLDTSTG-	121		
QY	60	PAACFPFGCTSLSTENOTLLEAGKMKGEKIAVASAHVILGPTRIMQRSL-----GGRG	114		
		: :			
Db	122	---YSGMATGATFNKDLFLDRGALGHEHNSGVHIALGPAY----GPLYKAARGNN	173		
QY	115	FESTGEDPPLAGGAALLIRGIQTSGVYATIKHFLLCNDQEDRRM-----	158		
		: :			
Db	174	FEAFGSDPLOGIAAAATIKGLENNWACVKHFETIGNEODIYROPSSKVDPEYDATKE	233		
QY	159	MOVISYTERALEEYIALPOLVAROSCGAMTYANGINSCEANSNKPYTDGLMLREKKM	218		
		: :			
Db	234	SISANTPDRAHMEHLMPPADSIR-AGVGSAVCYCINRYNNTSYSCENSYMTINHLKKEELDF	292		

QY	219	DCIMSDMGVYSTEAAVAGDLEMPB-----PRFGELTKFVNSGKPFHIVID	276
Db	293	OCFVVSDBMAQMSGAYSAISGLDMSMPBELLGWNTGCKSYWGQNLTKAVNEVPIERLD	355
QY	271	QKAREVL-----QCVKKCAASG--VTENGGETTVNN-----TP	301
Db	353	DMATRIALAAVATNSFPKRDLPNFSPTTKEGNEFFVDKTSPIYVKNNHVVDSNDFTE	411
QY	302	ETFAALLRKVNGEGVYLLKNNENVPLS--KKKTLI-----VGNPK-----QAT	344
Db	413	DTAL-----KYAESIYLLKNEKNTLPISPNKVRKLLSGIAGGPPPKGYECSIDQSCVGL	465
QY	345	YGGSGAALRAV--VAVPEFOLSKQLETPPSYTGAVATTVPILLEGCLLPDGAQPMRW	402
Db	470	EGGNSSGSV-GIPKQVTPFEISA-----	493
QY	403	RVFNEPPTPNRQHIDELFTKTDMLVDVYHPRAADTWADMGEGTYADECTYELGLV	462
Db	494	---NARKKKMQFDYIRESF-----DLQVSTVASDAMHSIVV	522
QY	463	VCGTAKAVDDQLVVDNATKQVPGDAFFGSAATREBTGHNLYKKNYTKFKIEGSAFTT	522
Db	529	SAVSEEGY-----LIID-----GN-----	542
QY	523	LKGDIVPGHGSILRVGCKVIDDOAEIEKSYALAKEHDOVITICAGLNADMETEGADRSM	582
Db	543	RGD-----KNNVTL-----MHNS-----	555
QY	563	KLPGLVDQLADVAANPNVYVM-QGTGPEEMFLDATPAVIAQWYG--GNETGSIAD	639
Db	556	-----DMLIKAVENCANTVYVITSQGVDESEFADHPNVTATVWAGPLADRSQTALAN	609
QY	640	VVEGDYNSGKLSISFPRLODNNAFLNFRFEG-----RTLYGSDYVGYRYAYEFAPKDV	695
Db	610	ILFGAANSGHLPFTVAASNDYIPITYVNPNEPEDNDLAEHDLVDVRYPEEKNIER	665
QY	696	NEPFGHGLSYTTFEAFSLSVSH-----	717
Db	670	RYARGYGLSYNEVYVSNAKVSAAKVDEBLPQRLYLAEYSYNKTEIINN-EDAFFPSNA	729
QY	718	-----KDKR-----LSLSLV	728
Db	730	RRIOEFLPYLDLSNWLTKDGVEYEDPYSTEORTTPIOPGGGLGDNALMEVAYKVEVD	789
QY	729	KNTGSVPQAQVLAOLYVYVPLQAKINRVPYKELKGAFKAYELOPGETKATYIEQGEKYVAYF	788
Db	790	QNLGNSTDKFYPQLTKHPBEKGKFTETV-QLRGFEKVELSPGEKTYVEFELLRDLSVM-	847
QY	789	DEERDQWCVERKDEYIVVSDSSAAKDGVAL	818
Db	848	DTTRQSWIVESGTVEYALI-----GVAV	869
RESULT 5			
BGL1_SACE1			
ID	BGL1_SACE1	STANDARD;	PRT; 876 AA.
AC	P22506;		
DT	01-AUG-1991 (Rel. 19, Created)		
DT	01-AUG-1991 (Rel. 19, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Beta-glucosidase 1 precursor (EC 3.2.1.21) (gentiobiose) (Cellulobiose)		
DE	(Beta-D-glucoside glucosyl/drolase).		
GN	BGL1		
OS	Saccharomycopsis fibuligera (Yeast).		
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
OC	Saccharomycetales; Saccharomycopsidaceae; Saccharomycopsis.		
OX	NCBI_TaxID:944;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RA	MEDLINE-89133518; PubMed-3146949;		
RA	Machida M., Ohtsuki I., Fukui S., Yamashita I.;		
RT	"Nucleotide sequences of Saccharomycopsis fibuligera genes for		

RT extracellular beta-glucosidases as expressed in *Saccharomyces cerevisiae*." RT Appl. Environ. Microbiol. 54:3147-3155(1988). CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-glucose residues with release of beta-D-glucose. CC -1- PATHWAY: Cellulose degradation. CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES. CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation of the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: M22475; AAA34314.1; CC InterPro: IPR002772; GH_3C. CC InterPro: IPR001764; GH_3N. CC Pfam: PF00933; Glyco_hydro_3; 1. CC Pfam: PF01915; Glyco_hydro_3_C; 1. CC PRINTS: PR00133; GLHYDRLASE3. CC PROSITE: PS00775; GLYCOSYL-HYDROL_F3; 1. KM Hydrolyase; Glycosidase; Cellulose degradation; Glycoprotein; Signal. FT SIGNAL 1 17 FT CHAIN 1 17 FT ACET SITE 18 876 BY SIMILARITY. FT CARBOHYD 22 22 N-LINKED (GLCNAC. .) (POTENTIAL). FT CARBOHYD 75 75 N-LINKED (GLCNAC. .) (POTENTIAL). FT CARBOHYD 224 224 N-LINKED (GLCNAC. .) (POTENTIAL). FT CARBOHYD 267 267 N-LINKED (GLCNAC. .) (POTENTIAL). FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL). FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL). FT CARBOHYD 372 372 N-LINKED (GLCNAC. .) (POTENTIAL). FT CARBOHYD 389 389 N-LINKED (GLCNAC. .) (POTENTIAL). FT CARBOHYD 426 426 N-LINKED (GLCNAC. .) (POTENTIAL). FT CARBOHYD 544 544 N-LINKED (GLCNAC. .) (POTENTIAL). FT CARBOHYD 585 585 N-LINKED (GLCNAC. .) (POTENTIAL). FT CARBOHYD 739 739 N-LINKED (GLCNAC. .) (POTENTIAL). FT CARBOHYD 780 780 N-LINKED (GLCNAC. .) (POTENTIAL). FT CARBOHYD 790 790 N-LINKED (GLCNAC. .) (POTENTIAL). SO SEQUENCE 876 AA; 96231 MW; DBB4AB9BE4A32F1E CRC64;

Query Match 16.1%; Score 706; DB 1; Length 876; Best Local Similarity 24.1%; Pred. No. 6e-38; Matches 238; Conservative 116; Mismatches 276; Indels 356; Gaps 29;

DB 411 AL--KVAEESIVLLKNNNTLPISTPEKAKLLISGIAAGPDPIGYOCEDOSCTNGALFQG 468
OY 348 GGSAAALRA--YVAATPFDGLSKOLETPPSYVAGATTVPILGBOCLTPDGPAPGRVRVN 406
DB 469 WSGSGSVSPKYYQYTPPEIS----- 488
OY 407 EPPQTPNRQHIIDELFFKTDMLVDVYHPKADTWADMEGYTADBCYTELGLVYCGT 466
DB 489 -----YLARKNNKQFDYIRE-----SYDAQYTKVA 514
OY 467 AKAYVDQQLVVDNAATKQVPDPAFFGSATREGRINLVKNTYKFKIERGSAPTYLLKGD 526
DB 515 SDAML--STVVVAAS----- 528
OY 527 TIVPGHSLVAGGCKVDDQAEIEKSYALAKEHDQVIGAGLNADMETEGADAPSKLPG 586
DB 529 -----GEQYIIVDG-----NGDKNKNTLWN 549
OY 587 VLDQLIADVAANPNVYVW-OTGTPREPMFLDTPAVIOAWYG--GNETGNSIADVFG 643
DB 550 NGDKLIETVAENCAATVYVVTSTGQINFEGFADHPNVTALVWAGPLGDSGTIANILFG 609
OY 644 DYNPSGLSTSPFKRLODNAPLNFTEAGRT-----LYGEDVYVGYRYEFAADKVNPF 699
DB 610 KAPSGHLPPTIAKTDYPIETSPSSGEPDNLVENDLLVDYREKKNIEPRVAF 669
OY 700 GHGLSYTTPAFSNLSVSH----- 717
DB 670 GYGLSTVEYVSAKSAKVAKVDDELPERATYLSFESYQAKSKNPDAFADNLRVN 729
OY 718 -----KDKL-----SVSLSVKNTGSVPGAQ----- 738
DB 730 EYLTPYLDNVTLKGDNYEYRPGDYSTEQRTTPNQGGGLGNDALWEVAYNSTDKFVPG 789
OY 739 -----VAQIYVPLQAAKINRVPYKELGKFAKVELDGPGEKATITIEQEKVAAAYDEER 792
DB 790 NSTDKFPLQYLYKHPEGKEFETP-QLRGFEKVELSPGEKTYDLRLRLRDSLVW-DTTR 847
OY 793 DQMCVEKGDYEVIVSDSSAKDVAL 818
DB 848 QSMIVESGYTEALI-----GVAV 865

RESULT 6
BGX SALTY STANDARD; PRT; 765 AA.
ID BGX SALTY
AC Q56078;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Periplasmic beta-glucosidase precursor (EC 3.2.1.21) (Gentibiase) (Cellulobiose) (beta-D-glucoside glucosylhydrolase) (T-cell inhibitor).
GN BGX OR STM2166.
OS *Salmonella typhimurium*.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; *Salmonella*.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RA Matsui K.;
RL Submitted (AUG-1996) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN-LT2 / SSGC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P., Courtney L., Porciliak S., Ali J., Dante M., Du F., Hou S., Layman D., Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E., Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M., Waterston R., Wilson R.K.;
RA "Complete genome sequence of *Salmonella enterica* serovar *typhimurium* LT2."

RL Nature 413:852-856(2001).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-glucose residues with release of beta-D-glucose.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 CC EMBL: D86507; BAA13102.1; -
 CC EMBL: AE008796; AAL21070.1; -
 CC StyGene; SG100604; bg1x.
 CC InterPro; IPR002772; GH_3C.
 CC InterPro; IPR001764; GH_3N.
 CC Pfam; PF00933; Glyco_hydro.3; 1.
 CC Pfam; PF01915; Glyco_hydro.3.C; 1.
 CC PROSITE; PS00775; GLYCOSYL_HYDROL_F3; 1.
 CC Hydrolase; Glycosidase; Periplasmic; signal; Complete proteome.
 KW SIGNAL
 FT CHAIN 1 20 POTENTIAL.
 FT ACT_SITE 21 765 PERIPLASMIC BETA-GLUCOSIDASE.
 FT CONFLICT 287 287 BY SIMILARITY.
 FT CONFLICT 630 630 R -> L (IN REF. 1).
 FT CONFLICT 634 635 EA -> KP (IN REF. 1).
 FT CONFLICT 656 656 V -> F (IN REF. 1).
 FT CONFLICT 674 674 E -> K (IN REF. 1).
 FT CONFLICT 717 717 E -> K (IN REF. 1).
 FT SEQUENCE 765 AA; 83392 MW; C7C767E6CAC96786 CRC64;
 SO
 Query Match 14.7%; Score 646.5; DB 1; Length 765;
 Best local similarity 25.1%; Pred. No. 3.4e-34;
 Matches 222; Conservative 124; Mismatches 309; Indels 229; Gaps 27;
 QY 3 DIDVEALIKKLTAEKVDLAGIDEMHTKALPKHGVPSLFTDQPNVGRGTKEFGVY 59
 DB 35 DAEVTDLKKMTVEDEKIGQLRLISVGPDN--PKAIREM-IKDQVGA----TFNVTRQ 87
 QY 60 -----PACPRCGTSGTSPNQLLLEAGKMGKEAI 91
 DB 88 DIRMQDQVWALSRLKPIEFPAVDVHGQRTVPISLGLASSNLDVATGVNSYERNA 147
 QY 92 AKSAHVILPTIMORSPLGSGFESIGEDPFLAGLAALINGIQT-----GYQATI 145
 DB 148 DDGLNMTWAPMVDVSRPBRMGASSEGEGEDTYLISNGETMVKMGKSPADRYSVMTSY 207
 QY 146 KHFELCNDQEDRRMMVOSI-VTERALREITLAPFOIARVDSQPCAFMTATNGINGVSCSEN 204
 DB 208 KHFPAAYGAVGSGREYNTVDMSOSRLENDMPYKAGL-DAGSGAVVAVANLSLGTPTSQD 266
 QY 205 PKYLDGKLRKMGMDGLINSDW-----YGVYSPTE-----AVVAGDLEMPGRRR 251
 DB 267 SWLTKVDLRDEMGKGLTYSDDHAIKELIKHGADEBAVRAALAGVDMSH--ADEYI 324
 QY 252 GETLKEVNSGKPFIVHIDQAREVLO-----FVKCAASGVTEGGETVYNNTPETA 304
 DB 325 SKYLRGLKSGVTMAELDQATRHVNVKVMGLFNDPYSHLGRKSDPDVTNAESRLHR 384
 QY 305 ALLRKVNGEIVLLKNENNVLPRLSKKKTLIVGRNA-KQATYGGGSAALRAYVATPPD 363
 DB 385 KEAREVAVRESVULLKKNLETLPLKKSGLTAVVGRPLADSORVDVGSWSAA----- 433
 QY 364 GLSKOLETPSYVGAATVPRILGSOCLPRGAPGRMVRVFERPGTNRQHIIDLEFT 423
 DB 434 GVANO-----SVTLVAGIQNAVGDAKTL----- 457
 QY 424 KTDMLHVDVYHPRADRTWYADMEGTATDEDCYEELGLVWCSGAKAYVDOLVVDNATRO 483
 DB 458 -----YAKGANITNDKGI-----VDFLNIYEAVKI 483

QY 484 VPGDAFGSAPRETERGRINLVKGNTRYKKEIEFGSAPTYTLKGDTVPQHGSLRVGCKYI 543
 DB 484 DPR----- 486
 QY 544 DDQATIEKSVLAKAKHDDVITICAGINADMETEGADRSMLPGVLDLIADVAANPNV 603
 DB 487 SPQAMIDRAVQAQKADADVAVAVGSGMAHEASSRNTITTPQSORDLITALKTKRPLV 546
 QY 604 VVMQGTGFEEMFWLD-ATPAVIOAWYCGNETGSIADVPEGDVNPSCGLSFPKRLQDN 662
 DB 547 LVLMNGRPLALVKEQQAADILETFMFACTEGMNAADLVLPEDVNPSCGLPISFPRSGOI 606
 QY 663 PAFINFEDEAGTLYGE--DVYVGYRYEFADKVDNPEPPGSLSTYTPAFSNLSVS----- 716
 DB 607 PLYYS-HINTGPPYPERPKNTYS-RYPDEANGPL-YFPGYGLSYTFEWSVDVTLSSPTM 663
 QY 717 HKDKLSVLSKKNNGSPCAQVNOVLVYKPIQAKINPNVYELKGFPAKVELQEPETAVY 776
 DB 664 QRDGKVTASVEYNTGKREGATVIOVYLDV-TASMSRPVQLKGFEKTLTKPGRKTVS 722
 QY 777 ---IEDEKVAAYFDEERDQWCEKGDYEVIVS-DSSAKDG 815
 DB 723 PFIDIE-----ALKFNNQMKYDAEPKFNVLGVDARVAKG 760
 RESULT 7
 BLX_ECOLI
 ID BGX_ECOLI STANDARD; PRT; 765 AA.
 AC P33633;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Periplasmic beta-glucosidase precursor (EC 3.2.1.21) (Gentlobiose) (Cellulobiose) (Beta-D-glucoside glucohydrolase).
 GN BGX OR B2132.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID:562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / W3110;
 RA Yang M., Luo S., Goddard A., Reilly D., Henzel W., Bass S.;
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / BHB2600;
 RA Richterich P., Lakey N., Gryan G., Jaehn L., Mintz L., Robison K.;
 RL Church G.M.;
 RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE-9742617; Pubmed-9278503;
 RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.D.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-glucose residues with release of beta-D-glucose.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC

Query Match 13.5%; Score 593; DB 1; Length 825;
 Best Local Similarity 24.5%; Pred. No. 1.1e-30;
 Matches 209; Conservative 111; Mismatches 262; Indels 272; Gaps 26;

9 ILKLTIAEVNLLAGID-----HTKALPKHGVSLFTDGPNGVGTKEFPNGVPA 62
 63 LVDMSTAEKVNLTGTGASGPGSGNTGSPRLNISTICVODGPILSRAADLT----- 117
 63 CFPQSTISGTFNOLLEAGKMKGAFAKSAHVILGPTIN--MORSPLGRGFESEIG 121
 118 VFCGMASSSPKQOLLYDAVAIGSFRRKGDALIGPYGMGKAAAGRGMEGHDP 177
 122 PFLAGLGAALLIRIGISTGVQATIKHFLCNDQEDRRM-----MVOS 162
 178 PYLEGVLAUQTIGIOSGVSTAKHLIGNEHEFRFAKKDKHAKIDPQMENTSSLS 237
 163 IYERALEIYALPFOJAVNDSPGAFMTAYNGINGVSCSENKYIDGMILRKMGWDGLI 222
 238 EIDDRAMHEIYLPFAAVANG--VSSIMCSYNKLNLSHACONSULLNYLKEELGFGFV 296
 223 MSDWYGTSTTEAVVAGLDLEMPREFRGETLKFVNSKGFPIHVIDORAREVLOFVK 282
 297 MTDMGALYSGIDANAGLDMDPCENQYFGNLTAVLNGITLPODLDMATRIISAL-- 354
 283 CAASGV--TENG-----ETVYNNPTEPALLRKYN 314
 355 -IYSGVNDPQPNYMAQTELTGHEXFKOEGDIVLANKHVNRSDINRAVALRS--AVBG 412
 315 IYLLKNNVLPKSKKTLI-----VGPNAK--QATYHGGSSALRAVYA----- 358
 413 VYLLKNEHETLPGRKRYRISILGSAADDSKTSICIRGCGSGAIGTGYSGAGTESY 472
 359 -VTPRGLSKOLETPPSYVGAATYVPILEGQCLTPDGAQPMKRVENEPGTEPNRQHI 417
 473 FYTPADGI-----GARAQOEKISYEFIG----- 495
 418 DELFFKTMHLVDYHPRKADRYADMEGTATADEDCYELGLVYVCGAKAYVDOLVY 477
 496 -----DSW-----NOAAMDSALYA 510
 478 DNATKQVPGDAFFGATRETRGRINLVKNTYKFKIEFSAPTYTLKGDITVPGHSLRV 537
 511 DAIEVA-----NSVAGEIGDVG--GN-----YGDINLNL----- 540
 538 GGCKVTDQAEIEKVAALAKEDQYITCAGLNADMETEGADRSKMLPGVIDOLIADYAA 597
 541 -----WHNAV-----PLIKINSS 553
 598 ANPNVYVVGOTGPPEM--PWLD--ATPAVIAQWYGNENETGNSIADVVGDNIPSKLSLS 654
 554 INNNTIYVTSQOQIDLEPFIDENENVTAVISSYLGDFVLAVALGDEPNSEKLPFT 613
 655 FPKRLQDN--PAFLNRTGAGRTLYGEDVYVGYRYEFAFDNPFEGSLSTTFPAFNL 713
 614 IAKDVNDYIPVLEKVDVDPDKFTESIYDYRFDKDKNKRYRFEFGISLSTNSLSDI 673
 714 SVSHDKLSVLSKNTGNSVPGAGVADLYKPLDAAKINRVELKGFANVELDOPGFTK 773
 674 EQ-----TLQPSSENAEPANYSLEY-----QYKSNMPPSE-- 706
 774 AVTIEOEKYYAAV 787
 707 -YTVPEGEKELANY 719

RESULT 9
 BGLI ASPAC STANDARD; PRT; 860 AA.
 AC P48825;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 13-JUN-2002 (Rel. 41, Last annotation update)

DE Beta-glucosidase 1 precursor (EC 3.2.1.21) (Gentiobiase) (Cellbiase)
 DE (Beta-D-glucoside glucosylhydrolase).
 OS Aspergillus aculeatus.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus;
 NCBI_TaxID=5053;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=F-50;
 RX MEDLINE=97082984; PubMed=8964516;
 RA Kawaguchi T., Enoki T., Tsurumaki S., Sumitani J., Ueda M.,
 RA Ooi T., Arai M.;
 RT "Cloning and sequencing of the cDNA encoding beta-glucosidase 1 from
 RT Aspergillus aculeatus";
 RL Gene 173:287-288(1996).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
 CC glucose residues with release of beta-D-glucose.
 CC -1- PATHWAY: Cellulose degradation.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL: D64088; BAI10968.1; -
 DR InterPro: IPR002772; GH_3C.
 DR InterPro: IPR001764; GH_3N.
 DR Pfam: PF00933; Glyco_hydro_3; 1.
 DR Pfam: PF01915; Glyco_hydro_3_C; 1.
 DR PRINTS: PR00133; GLYHDLASE3
 DR Prosite: PS00775; GLYCOSYL_HYDROL_3; 1.
 KW Hydrolyase; Glycosidase; Cellulose degradation; glycoprotein; signal.
 FT SIGNAL 1
 FT CHAIN 20
 FT ACT_SITE 280
 FT CARBOHYD 61
 FT CARBOHYD 211
 FT CARBOHYD 252
 FT CARBOHYD 315
 FT CARBOHYD 322
 FT CARBOHYD 354
 FT CARBOHYD 387
 FT CARBOHYD 442
 FT CARBOHYD 523
 FT CARBOHYD 542
 FT CARBOHYD 564
 FT CARBOHYD 658
 FT CARBOHYD 668
 FT CARBOHYD 690
 FT CARBOHYD 712
 SQ SEQUENCE 860 AA; 93052 MW; 4B464778B00FC694 CRC64;

Query Match 13.2%; Score 581; DB 1; Length 860;
 Best Local Similarity 23.3%; Pred. No. 7e-30;
 Matches 225; Conservative 114; Mismatches 299; Indels 326; Gaps 28;

8 ALLKLTIAEVNLLAGIDFW-----HTKALPKHGVSLFTDGPNGVGTKEFPNGV 60
 49 AIVSQMTLDEKVNLTGTG--WELEKCVGTGGVPRNLNIGMCLDPSLDIRSDY----- 102
 61 AACFPQSTISGTFNOLLEAGKMKGAFAKSAHVILGPTIN--MORSPLGRGFESEIG 119
 103 NSAFPAGVAVNATWKNLAVYLRQAGMGQFSDKIDVQLPAPAGPLGNSPDGGRWMEGFS 162
 120 EDPFLAGLGAALLIRIGISTGVQATIKHFLCNDQEDRRM-----VOSIYTE 166
 163 PDEALGVLFATIKINGIDQAGVATKHYILNDQEHFRQVABAAGYGNISDSSNVD 222
 167 RALREIYALPFOJAVNDSPGAFMTAYNGINGVSCSENKYIDGMILRKMGWDGLMSDW 226

[illegible]

RA MEDLINE=90175009; PubMed=2106673;
 RA Ohmura K., Takano M., Shimizu S.;
 RT "DNA sequence of a beta-glucosidase from *Ruminococcus albus*."
 RL Nucleic Acids Res. 18:671-671(1990).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
 CC glucose residues with release of beta-D-glucose.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X15415; CA3461.1; -
 DR PIR: S08243; S08243.
 DR InterPro: IPR001764; GH_3N.
 DR Pfam: PF00933; Glyco_hydro_3; 1.
 DR Pfam: PF01915; Glyco_hydro_3_C; 1.
 DR PRINTS: PR00133; GLHYDRLAS3.
 DR PROSITE: PS00775; GLYCOSYL_HYDROL_F3; 1.
 DR Hydroxylase: Glycosidase; Cellulose degradation.
 FT ACT_SITE 696 696
 FT ACT_SITE 947 AA; 104277 MW; C43B8CDD9D60A115 CRC64;
 SQ
 Query Match 9.7%; Score 425; DB 1; Length 947;
 Best Local Similarity 34.7%; Pred. No. 9.6e-20;
 Matches 107; Conservative 50; Mismatches 105; Indels 46; Gaps 6;
 QY 31 KALPHGVPSLRTFDGNGVR---GYFENGVPAPCCPGCTSLGTFNQTLLAEAGKMMG 87
 DB 505 KHELEGLIPACCSDDGSGMRLDVGTRAFS-----LPNGTLAATFNKSLTELEPTIYG 558
 QY 88 KEATKSAHVILGPTLNNORSPGGRGESIGDEPPLAGGAALLRGIOSTVOATYTH 147
 DB 559 LEMANRNVKDCILGPGGMNTHHPLNGRFEYFSEDPFLTGMAALEGLHSGVEGTIKH 618
 QY 148 FLCDQEDRRMVOISVTEREYALPPQIAVRDSOPGAFATVANGINGVSCSENPXY 207
 DB 619 FCANNOETNRHFIIDSVASERALEIKLGEIYAVRSKASVMTYTKVNGMLTWSFPL 678
 QY 208 LDGMKREMGWDLGIMSDWYGTST-----TEAVVAGLDLEMPGPPRRGTELKE 257
 DB 679 NTMLRKONGFDTMTDMANINDRGACAPDKNNFAAMVRAQNDVYMCAD---GEGSD 735
 QY 258 NV---SNGKPFIVHIDQAREVLYQFKKCAASGVTEENGPEPTTVNTPETALLRKVME 313
 DB 736 NVIALADGRUTRAELQRSARNILSFM-----STHMARKLGED 775
 QY 314 GIVLTKKE 321
 DB 776 EAVEVINK 783.
 RESULT 12
 YBBD_BACSU
 ID YBBD_BACSU STANDARD; PRT; 642 AA.
 AC P40406;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical lipoprotein ybbd precursor (ORF1).
 GN YBBD.
 OS *Bacillus subtilis*.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_Taxid=1423;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=BD99 / MS94;

RX MEDLINE=94281248; PubMed=8011666;
 RA Quirk P.G., Guffanti A.A., Clejan S., Cheng J., Krulwich T.A.;
 RT "Isolation of Tn917 insertion mutants of *Bacillus subtilis* that are
 RT resistant to the protonophore carbonyl cyanide
 RT m-chlorophenylhydrazine."
 RL Biochim. Biophys. Acta 1186:27-34(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168.
 RX MEDLINE=97419514; PubMed=9274029;
 RA Liu H., Haga K., Yasumoto K., Ohashi Y., Yoshikawa H., Takahashi H.;
 RT "Sequence and analysis of a 31 kb segment of the *Bacillus subtilis*
 RT chromosome in the area of the *trnH* and *trnG* operons."
 RL Microbiology 143:2763-2767(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunet F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerion I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Colglinty E.J., Grandi A.,
 RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holstappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Katamata D., Kishihara Y., Kjaer-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kunita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestel D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Pressac A., Puig P., Purnelle B., Rapoport G., Ray M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sedate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeder R., Scofield F.,
 RA Seliguchi J., Sekowski A., Seror S.J., Serior P., Shin B.S., Soldo B.,
 RA Sorokin A., Taccoti E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandendol M., Vannier F., Vassaretto A.,
 RA Viali R., Wambuit R., Wedler E., Wedler H., Wellenreger T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
 RT *subtilis*."
 RL Nature 390:249-256(1997).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (potential).
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L19954; AAA64351.1; -
 DR EMBL: AB002150; BAA19499.1; -
 DR EMBL: 299104; CAB11942.1; -
 DR Subtilist; BG10832; YbBD.
 DR InterPro: IPR001764; GH_3N.
 DR InterPro: IPR001764; GH_3C.
 DR Pfam: PF00933; Glyco_hydro_3; 1.
 DR Pfam: PF01915; Glyco_hydro_3_C; 1.
 DR PROSITE: PS00775; GLYCOSYL_HYDROL_F3; 1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Hypothetical protein; Hydrolase; Glycosidase; Membrane; Lipoprotein;
 KW Signal; Complete proteome.

FT SIGNAL 1 16 POTENTIAL.
 FT CHAIN 17 642 HYPOTHEICAL LIPOPROTEIN YBBD.
 FT LIPID 17 17 N-ACYL DIGLYCERIDE (POTENTIAL).
 FT ACT_SITE 318 318 BY SIMILARITY.
 SQ SEQUENCE 642 AA; 70580 MW; DCEA93142922P13F CRC64;
 Query Match 5.0%; Score 219.5; DB 1; Length 642;
 Best Local Similarity 21.5%; Pred. No. 1e-06;
 Matches 101; Conservative 75; Mismatches 126; Indels 167; Gaps 20;
 QY 4 IDVEALIKKLTLEAKVDLAGITF--W-----
 DB 33 IDANQIVNRSLEDEKIGLMPDFRMWOKGESSPOLATKMNDEVASLYKKYFGGIIIF 92
 QY 29 -----HTKLPKHGVPSLRFETDGNV-----RGKRFNGVPAACPG 67
 DB 93 AENKTKKQTVLTDIYQKASPK--IPLMISIDGEGIVRLGEGIN-----PFGN 141
 QY 68 TSLGTFNQTLEBAGKMGKKAIAKSAHVILPTINMRSP---LGGRGFESIGEDPF 123
 DB 142 MALGAANSRLNAVQTSIIIGKELSAIGINDFSPVYDINNPNPIYGVRSFSNELTS 201
 QY 124 LAGIGAAALRGIOSTGVQATIKFLCNDQEDRRMAYQS-----IYT--ERALREIYALP 176
 DB 202 RLGL--YTKKGLQRODIAASALKHFPCHGDTD---VDSHYGLPLVSHGGERLREVELY 254
 QY 177 FOIARNDGOGAPFATY-----NGING-----VSCSENPYLDGMLRKGMDP 220
 DB 255 FQKAI-DAGADWMTAVTAVQPPADDTYKSKLDGSDILVATSKVYGLLEQMGFNG 313
 QY 221 LIMSWMYGT-----YSTEAVY---AGLDLMPGPPRPGETLKNVNSNGKPFTHVID 270
 DB 314 VIVTDALNMKAIAIDHFGOEAEVAVYAKAGVIAL-----MPASVSLKE-----E 358
 QY 271 QRAREVQIPKCKAASGVTEGPEPTVYNNPERRALL----- 307
 DB 359 QKFAVIAOLKAEVKNK--DIPQOQINNSVERTISLTKIRGMYPARNSDSTEKIAKAK 415
 QY 308 -----RKVNEGIVLTKENNVPLSKK--KTLIVGPNAKQ 342
 DB 416 KIYSGKHLEAKKLAKEKATVLEKNEQHTLPFKPKGSRLIYAPYEQ 464
 RESULT 13
 BGLS SCHCO STANDARD; PRT; 192 AA.
 ID BGLS SCHCO STANDARD; PRT; 192 AA.
 AC P29091;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE beta-glucosidase (EC 3.2.1.21) (Gentlobase) (Celllobase)
 DE (beta-D-glucoside glucosylhydrolase) (Fragment).
 OS Schizophyllum commune (Bracket fungus).
 OS Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
 OC Agaricales; Schizophyllaceae; Schizophyllum.
 OX NCBI_TaxID=5334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=66295810; PubMed=3091028;
 RA Morandell M., Barbier J.R., Dove M.J., Mackay R.M., Sellig V.L.,
 RA Yaguchi M., Willik G.E.;
 RT "A clone coding for Schizophyllum commune beta-glucosidase: homology
 RT with a yeast beta-glucosidase";
 RL Biochem. Int. 12:905-912(1986).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
 CC glucose residues with release of beta-D-glucose.
 CC -1- PATHWAY: Cellulose degradation.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M27313; AAA33925.1; -.
 DR PIR: A28571; A28571.
 DR InterPro: IPR002772; GH_3C.
 DR InterPro: IPR001764; GH_3N.
 DR Pfam: PF01915; Glyco_hydro_3_C; 1.
 DR PROSITE: PS00775; GLYCOSYL_HYDROL_F3; PARTIAL.
 KW hydrolase; glycosidase; cellulose degradation; glycoprotein.
 FT NON_TER 1
 SQ SEQUENCE 192 AA; 20396 MW; 2B410C1805E648A5 CRC64;
 Query Match 3.7%; Score 164.5; DB 1; Length 192;
 Best Local Similarity 34.5%; Pred. No. 0.00066;
 Matches 40; Conservative 19; Mismatches 52; Indels 5; Gaps 3;
 QY 589 DDLIADVAANPNVYVMOT-GTPEMPWLDATPAVIAWYG--GNETGNSIADVFGDY 645
 DB 77 DALVQAVADANENTIVAANTVCAITTEAWIEHPNKAVYVSGLPQEGANSVADILYGAY 136
 QY 646 NNSGKLSLFPRLQDNPAFLNRTFA--GRLYGEDVYVGYRYRFPDKVNPF 699
 DB 137 NPSGRLPYIAKSADYPAQVLYESSAQVPIIDYSEGLVDRHFDANGIEPRFE 192
 RESULT 14
 MAG2_VIBCH STANDARD; PRT; 330 AA.
 ID MAG2_VIBCH STANDARD; PRT; 330 AA.
 AC Q9K037;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE beta-hexosaminidase (EC 3.2.1.52) (N-acetyl-beta-glucosaminidase)
 DE (Beta-N-acetylhexosaminidase).
 GN MAG2 OR VC0692.
 OS Vibrrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=El Tor N16961 / Serotype O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eissen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Dragoi I., Sellers P.,
 RA Ermolaeva M.D., Yamathyan U., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Ullrich T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*
 RT cholerae";
 RL Nature 406:477-483(2000).
 CC -1- FUNCTION: Cleaves GlcNAc linked beta-1,4 to MURNAc tripeptides
 CC (by similarity).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal non-reducing N-
 CC acetyl-D-hexosamine residues in N-acetyl-beta-D-hexosaminides.
 CC -1- PATHWAY: CELL WALL SYNTHESIS; MUREIN TRIPEPTIDE RECYCLING PATHWAY.
 CC -1- SUBUNIT: MONOMER (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES. MAG2
 CC SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: AE004155; AAF93857.1; -.

Query Match	3.5%;	Score 151.5;	DB 1;	Length 330;
Best Local Similarity	22.8%;	Pred. No. 0.0099;		
Matches 75;	Conservative 44;	Mismatches 11;	Indels 99;	Gaps 12;

Qy	8	AIKIKLITAEVVDLACIDEFNHTKALRKHGVSLSRFTDGGVGRGTFFNGVAPCPCG	67
Dd	44	ALNKAIQNAKRPFLILICVD-----QBGGRVGRFREG-----FERIPAQYAR	86
Qy	68	TSLGSTFQOTLLEAGKMKGKEAIKASANYILPTIM--QRSPDLGRGFEESGEDPFLA	125
Dd	87	AENG-----VELAEGGWMMAELTAHDVDSLFAFVLDLGPACKAIGNRAF--GEVQTV	139
Qy	126	GIGAAALIRIGIOSTGVQATIKHF-----LCNDQEDRIMVYOSITERALREIYA	174
Dd	140	LKHSASFALRGKKAVMATTFGHHFGHGAVALADSHLETTPYDERETIINQDMAIFQAIGAGV	199
Qy	175	LRFQIARV-----DSQPGAFMYALNGINGVSCSENNKIYDGMIRKEMGMDGLMSDWIG	228
Dd	200	LDAMMPAHVYVPHYADP-----ASGSYMWIKOVALRETLGPKGIVFSD--	242
Qy	229	TYSTTEAVAGLIDLEMGPREFRGETLKFVNSMGKPIIHVIDRARE-----VDFV	280
Dd	243	-----DLMSGAAVMGSPV-----BRSHALVAGGDMILICN	274
Qy	281	KKCAASGVTEGPEPTYNNNPETRALLRK	309
Dd	275	KREAAVEVDNLV---IMEVPOAEALLRK	300

	RESULT	15
ID	NAGZ_VIBPU	
AC	P66157;	
DT	15-JUL-1998 (Rel. 36, Created)	
DT	30-MAY-2000 (Rel. 39, Last sequence update)	
DT	16-OCT-2001 (Rel. 40, Last annotation update)	
DE	Beta-hexosaminidase (EC 3.2.1.52) (N-acetyl-beta-glucosaminidase)	
DE	(Beta-N-acetylhexasaminidase).	
GN	NAGZ OR EXO II.	
OS	Vibrio furnissii.	
OC	Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.	
OX	NCBI_TaxID=29494;	
RN	[1]	
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 1-16.	
RC	STRAIN=7225;	
RA	MEDLINE=97125984; PubMed=8969206;	
RX	Chittaru E., Roseman S.;	
RT	"Molecular cloning and characterization of a novel beta-N-acetyl-D-	
RL	glucosaminidase from Vibrio furnissii.";	
RN	J. Biol. Chem. 271:33433-33439(1996).	
RP	[12]	
RP	REVISIONS TO 70-80.	
RU	Chittaru E., Roseman S.;	
RA	Submitted (FEB-1999) to the EMBL/Genbank/DDBJ databases.	
CC	-I- FUNCTION: Cleaves GlcNAc linked beta-1,4 to murine tripeptides	
CC	(By similarity). HYDROLYZES RAPIDLY P-NITROPHENYL-N-ACETYL-BETA-D-	
CC	GLUCOSAMINIDE (PNP-BETA-GLCNAc) AND 4-METHYLBELLIFERYL-BETA-	
CC	GLCNAc, AND SLIGHTLY ACTIVE ON p-NITROPHENYL-BETA-GALNAc. MAY PLAY	
CC	A ROLE IN SIGNAL TRANSDUCTION BETWEEN HOST AND ORGANISM.	
CC	-I- CATALYTIC ACTIVITY: Hydrolysis of terminal non-reducing N-	
CC	acetyl-D-hexosamine residues in N-acetyl-beta-D-hexosamides.	
CC	-I- ENZYME REGULATION: INHIBITED BY GLCNAc, 2-ACETAMIDO-1-N-(4-L-	
CC	ASPARTATE)-2-DEOXY-BETA-D-GLUCOPYRANOSYLAMINE (GLCNAc-ASN) AND O-	
CC	(2-ACETAMIDO-2-DEOXY-D-GLUCOPYRANOSYLIDENE)-AMINO-N-	

```

CC      PHENYLCARBAMATE (PUGNAC).
CC      -I- PATHWAY: CELL WALL SYNTHESIS; MUREIN TRIPEPTIDE RECYCLING PATHWAY.
CC      -I- SUBUNIT: MONOMER (POTENTIAL).
CC      -I- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC      -I- MISCELLANEOUS: MAXIMAL pH AND OPTIMAL TEMPERATURE OF THE ENZYME
CC      ARE 7.0 AND 45 DEGREES CELSIUS.
CC      -I- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES. NAG2
CC      SUBFAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC      or send an email to license@isb-slb.ch).
CC      -----
DR      EMBL; U52818; AAC44686.2; -.
DR      InterPro; IPR001764; GR_3N.
DR      Pfam; PF00933; Glyco_hydro_3; 1.
DR      PROSITE; PS00775; GLYCOSYL_HYDROL_F3_1.
KW      Hydrolase; Glycosidase; peptidoglycan synthesis; Cell division;
KW      Cell wall.
KW      ACT_SITE
FT      242
SQ      SEQUENCE 329 AA; 36181 MW; 440C6A9B18143C34 CRC64;
           BY SIMILARITY

```

Query Match	3.4%;	Score 147.5;	DB 1;	Length 329;
Best Local Similarity	23.3%;	Pred. No. 0.018;		
Matches	73;	Conservative	43;	Mismatches 114; Indels 83; Gaps

QY	16	AERVDLLAGIDFWHMKLPKRGVPSLRFDTDGPNGVKGTRFFENGVPACPGCTSIGSTFN	75
	52	AAKRPILIGVD-----GGGRVORFRDG-----FSKIPAAQLYARSDNGT---	91
QY	76	QTLLPEAGKMKCKEALAKSAHVILGPITN--MQRSPGGRGREGSIEGDEPFLAGLGAALI	137
QY	92	-QIAEKGGMVMAAEILIAHDIDLSPAPVLADGFCFCRAIGNRAF--GDVQTVLYTSSAYM	147
QY	134	RGIOSTPGVQATIKHP-----LCNQDEGRMMVQVSIYTERALREYALPOTAVR	182
Db	148	RKMSVGMATTKKHPGCGAVIADSHLETYPDERDSIADDMITFRQIETAGILDMMPAH	207
QY	183	-----DSOPGAFMTAYNGINGVSCSENPXYLDGMRLKMGDGLMSDMGTGYSTTEAV	236
Db	208	VIVPHVDAP-----ASSPFWLQVLRQELGEGIVFSF-----	242
QY	237	VAGLDLMPGPRFENGTELKKNVNGKPRFIHVIDQAREYLDQPVKCAASGYTEMGPETT	296
Db	243	-----DLSMGAAIMGSPERAQDS-----LDACGDVLMCNKRRESVAVLDOLPISV	290
QY	297	VNNTPETALLRK 309	
Db	291	V---PQAQSLKQ 300	

Search completed: April 26, 2003, 13:08:55
Job time : 36 secs